STAT 530/J530
August 23rd, 2005

Instructor: Brian Habing
Department of Statistics
LeConte 203
Telephone: 803-777-3578
E-mail: habing@stat.sc.edu

Introduction to R

R can be downloaded for free from www.r-project.org

This information and all of the code used today can be found at:
http://www.stat.sc.edu/~habing/courses/530rbigF05.html

Saving Your Work

All of the data you create in a session can be saved when you quit

- Automatic if you have administrative access
- Use **`Save Workspace`** and **`Load Workspace`** (in the **`.Rdata`** file) otherwise.

It only saves the objects you created!

Some Important Commands

- **`objects()`**
- **`library(MASS)`**
- **`help(`***function name*`)`**
- **`( )`** indicate a mathematical grouping or arguments of a function
- **`[ ]`** indicate you are finding an element of a vector or matrix

Basic Example

```
x<-c(5,4,3,2)
x[3]
mean(x)
objects()
hist(x)
t.test(x,alt="greater",mu=5)
x.t<-t.test(x,alt="greater",mu=5)
attributes(x.t)
x.t$p.value
```


Writing Your Own Function

```
weird<-function(x){
  y<-sin(x)+x^2
  y}
weird(5)
weird(x)
weird<-function(x=7){
  y<-sin(x)+x^2
  y}
weird(x)
weird()
```

Using a Large Data Set

The following code looks at the **oildata** data set from last class.

```
oildata<-
   read.table("http://www.stat.sc.edu/~habing/
   courses/data/oildata.txt",header=TRUE)
oildata
attributes(oildata)
```


oildata (continued)

```
oildata[,2]
oildata[2,]
oildata$Gender

oildata[oildata[,2]=='1',
  c("Marital","Age","Econ","Conv",
  "Q1")]
```


oildata (continued)

```
apply(oildata[,12:31],1,sum)
tapply(oildata$Low,oildata$Gender
  ,mean)
plot(oildata$Income,oildata$Low)
plot(as.factor(oildata$Income),
     oildata$Low)
plot(as.factor(oildata$Income),
      oildata$Q1)
```

More Complicated Example

Determine the total score for each person on Q1 to Q20 (reversing any questions that seem to need it).

Then make side-by-side boxplots of the total for each age group, with a dotted line for the overall average, and the years for the different age group showing on the x-axis.

The Code For It

```
corvals<-cor(oildata[,12:31],oildata$Low)
posones<-(1:20)[corvals>0]
negones<-(1:20)[corvals<0]
sumqs<-apply(oildata[,11+posones],1,sum)
  + apply(6-oildata[,11+negones],1,sum)
plot(as.factor(oildata$Age),sumqs,
  xaxt="n")
axis(side=1,at=1:5,labels=c("18-25","26-
  35","36-45","46-65",">65"))
lines(c(0,6),c(mean(sumqs),mean(sumqs)),
  lty=3)
```


Homework

Using R, make two variables containing the ratings of "Low Energy Use", one for males and one for females.

Conduct a two sample t-test to see whether there is a difference between the genders and check the assumptions using a q-q plot. Summarize your results.

Hint: `t.test(x,y)qqnorm(x)qqline(x)`